Chemical Education

A CHIMIA Column

Topics for Teaching: Chemistry in Nature

Protecting the eggs of a praying mantis: natural biomaterials[a]

Catherine E. Housecroft*

*Correspondence: Prof. C.E. Housecroft, Department of Chemistry, University of Basel, BPR 1096, Mattenstrasse 24a, CH-4058 Basel, Email: catherine.housecroft@unibas.ch

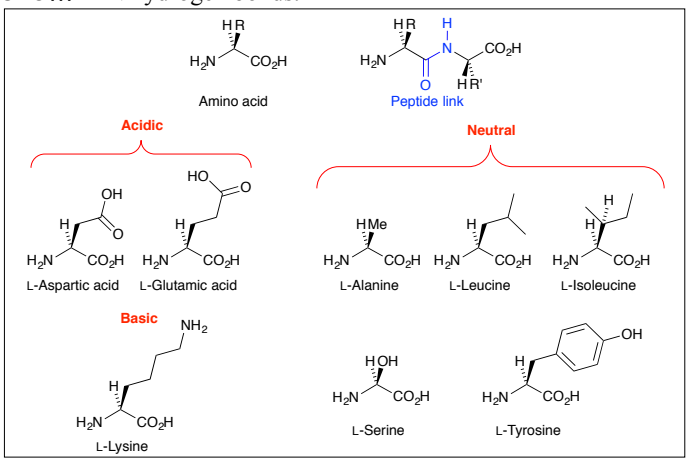
Abstract: Coiled-coil proteins are the basis of the biomaterial that protects the overwintering eggs of praying mantises.

Keywords: Praying mantis · Biomaterial · Protein · Coiled-coil Ootheca · Chemical education

Praying mantis is the common name for members of the family of carnivorous insects called *Mantidae* which belong to the order *Mantodea*. In southern Europe, *Mantis religiosa* (genus *Mantis*) is commonly encountered (Fig. 1), and its geographical range extends through Africa, Asia, Australia and North America. It is a relatively large mantis which, in Europe, has a life cycle of one generation per year, overwintering at the egg stage. The eggs are enclosed within a protective case called an *ootheca* which can withstand snow and very cold winter temperatures. The ootheca in Fig. 2 was found at an altitude of 1200 m in south-eastern France. Not all mantids overwinter as eggs. In southern Europe, eggs of the conehead mantis, *Empusa pennata*, hatch in late summer, spending the winter as nymphs and reaching the adult stage in early summer. A female mantis lays her eggs within a white foam-like mass that hardens into a distinctive shape (Fig. 2). Depending upon the genus of mantis, an ootheca may contain between 15 to 80 eggs and early development of the young insects takes place within the egg-case. An ootheca is composed largely of protein [1,2], and before continuing, we review some terms concerning protein structure.

Fig. 1. The praying mantis *Mantis religiosa*. ©Edwin C. Constable 2018

Nature is equipped with a pool of amino acids, some of which are shown in Scheme 1. A polypeptide comprises a sequence of amino acids coupled through peptide links (Scheme 1), and a protein is a high molecular mass polypeptide. The sequence of amino acids in a protein defines its *primary structure*. Functionalities such as a carboxylic acid, amide, thiol or aryl group that can participate in hydrogen bonds, disulfide bridges (see an earlier *Chemical Education Column* [3]), ionic interactions between charged residues, and hydrophobic contacts. Such interactions in a protein lead to its *secondary structure*: *α-helices, β-sheets, turns* and *coils*, and to the *tertiary structure* which is the overall protein conformation. In a polypeptide chain with an *α*-helix, the conformation is stabilized by C=O…H–N hydrogen bonds.



Scheme 1. General structure of an amino acid, and condensation of two amino acids to form a peptide link. A dipeptide (and any polypeptide) has *N*- and *C*-terminii (at the left and right in the diagram, respectively). Examples of naturally occurring amino acids with acidic, basic and neutral groups.

Fig. 2. An ootheca (egg case) of a mantis. ©Edwin C. Constable 2018.

The protein-containing sheets that make up a mantis ootheca comprise a *coiled-coil* tertiary structure. Coiled-coils were independently first recognized in 1953 by Francis Crick (who coined the phrase coiled-coil) and Linus Pauling (who used the term 'compound helices') [4]. They are formed by proteins with a repeating seven amino acid residue sequence (a *heptad*) [5]. This sequence is usually represented as $(abcdefg)_n$ and the amino acids in positions *a* and *d* are more hydrophobic than those in other positions. Hydrophilic residues occupy positions *e* and *g*. Two (Fig. 3) or more *α*-helices combine together into a supercoiled structure, the key to assembly being the respective positions of the hydrophobic and hydrophilic domains and the interactions between them. Scheme 2 shows a

'wheel' representation of a dimeric coiled-coil with parallel orientations of the polypeptide chains. Interactions between hydrophilic residues give rise to salt bridges, while those between hydrophobic residues (*a* and *d* in Scheme 2) produce so-called 'knobs-into-holes' packing [5]. The coiled-coil structure depends upon the hydrophobic and hydrophilic character of an amino acid rather than its specific identity. As a result, a large variety of amino acid sequences can be accommodated in a coiled-coil structure. In helices making up the coiled-coil assembly, the hydrophobic *a* and *d* residues reside within the coil (Scheme 2) and are shielded from the solvent which surrounds the protein in solution.

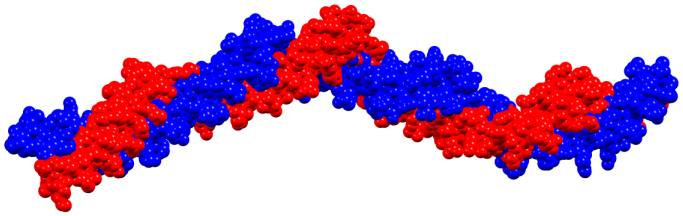
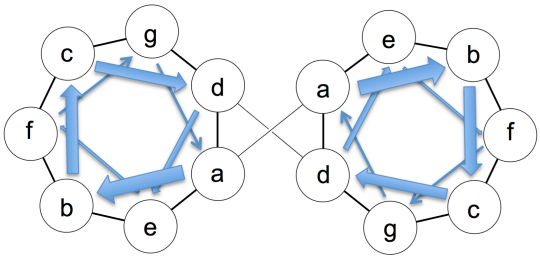
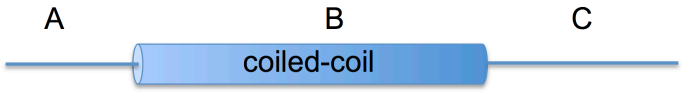


Fig. 3. Representation of a coiled-coil protein containing two strands (PDB code 1C1G [6]).



Scheme 2. Heptad wheel representation of a coiled-coil dimer with the helical polypeptide chains in a parallel orientation. The arrows show the direction of the helical twist around each *abcdefg*-unit which repeats along the chain.

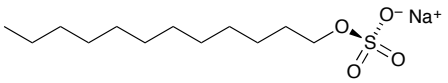
Coiled-coils are ubiquitous in Nature [4,5] and >90% consist of dimers (Fig. 3), trimers or tetramers for which the algorithms for choice of amino acid and the interactions between them are well understood [7]. The hydrophobic positions of the heptads are predominantly occupied by the amino acids L-leucine and L-isoleucine. Mantis ootheca and silks from some *Hymenoptera* (see below) are unusual in having L-alanine as the dominant hydrophobic amino acid. L-Alanine is less hydrophobic than L-leucine and L-isoleucine, and consequently, higher concentrations of the ootheca-forming proteins are required for protein folding. The required protein concentrations are reached by production and storage in a specialized gland in a female mantis. Highly concentrated aqueous solutions of protein in the form of aerated foams are secreted during egg-laying. Larvae of some bees and ants (which belong to the order *Hymenoptera*) produce silks comprising tetrameric coiled-coils. The latter contrast with the protein silks of spiders and silk-moth larvae which feature *β*-sheets. Cocoon and nest silks of bees and ants are extremely light but are very tough. Honey-bee larvae produce protein silks to strengthen the wax cells in which they pupate [8]. The general form of mantis dimeric coiled-coils, and bee and ant tetrameric coiled-coils is that of a *triblock macromolecule* represented in Scheme 3. The lengths of the blocks A, B and C vary with the insect, but the central block consists of the coiled-coil protein section with around 200 amino acid residues but no primary sequence (i.e. heptads are present with variable compositions, see above). The A and C blocks comprise the *N*- and *C*-terminal regions (see Scheme 1) with a primary sequence and *β*-sheets comprising the secondary structure.



Scheme 3. Mantis and Hymenoptera silks are triblock (ABC) macromolecules in which blocks A and C are the *N*- and *C*-terminal regions of the protein.

The high L-alanine content in mantis ootheca and silks from the larvae of some bees and ants is curious. Although this amino acid is highly favourable for *α*-helix assembly, its weak hydrophobicity leads to less strong 'knobs-into-holes' packing within the coiled-coil. This seems to be compensated for by the particularly long B-blocks. Another advantage of L-alanine is the small side chain (Scheme 1) which is probably advantageous for crystallite formation [9]. Indeed, the protein silks in mantis ootheca are highly crystalline with an ordered structure following the fibre axis [10]. Crystalline polymers as found in mantis ootheca are stronger and tougher than non-crystalline polymers. Another peculiar feature of mantis ootheca proteins is the replacement of L-alanine by L-tyrosine (Scheme 1) in a significant number of *d* positions in the heptad (Scheme 2) [2]. This introduces aromatic amino acids which are more typically associated with *β*-sheet-containing silks [11].

Once a mantis ootheca hardens (Fig. 2), its analysis by physical and chemical means is difficult. Thus, for the purpose of chemical analysis, material has been collected during egg-laying [2]. The number of proteins present and their molecular masses have been determined by using the analytical technique of SDS-PAGE (an acronym for *sodium dodecyl sulfate–polyacrylamide gel electrophoresis*). The protein samples are first treated with sodium dodecyl sulfate (Scheme 4) which denatures the proteins by destroying their secondary and teriary structures. This process generates unfolded polypeptide chains which associate with the dodecyl sulfate anions. The size of the protein determines the amount of dodecyl sulfate with which it associates, and this in turn determines the charge of the protein-dodecyl sulfate unit. During gel electrophoresis, different sized protein-dodecyl sulfate units are separated on the basis of their charges [12]. Protein composition and structure have been investigated using solid-state ^{13}C NMR and IR spectroscopies [1,2] and electron diffraction [10].



Scheme 4. Structure of sodium dodecyl sulfate.

This column has highlighted Nature's coiled-coil proteins with particular reference to the role they play in protecting the overwintering eggs of praying mantises. Nature inspires many aspects of science, and coiled-coils are important for engineered biomaterials, either in the form of *regenerated* silk proteins (i.e. natural) or *recombinant* silk proteins (i.e. artificially produced) [13].

References

[1]. K.J. Kramer, V. Bork, J. Schaefer, T.D. Morgan, T.L. Hopkins, *Insect Biochem.* **1989**, *19*, 69.
[2]. A.A. Walker, S. Weisman, T. Kameda, T.D. Sutherland, *Biomacromolecules* **2012**, *13*, 4264.
[3]. C.E. Housecroft, *Chimia* **2018**, *72*, 428.
[4]. A.N. Lupas, J. Bassler, *Trends Biochem. Sci.* **2017**, *42*, 130.
[5]. F. Lapenta, J. Aupič, Ž. Strmšek, R. Jerala, *Chem. Soc. Rev.* **2018**, *47*, 3530.
[6]. The Protein Data Bank: www.wwpdb.org
[7]. T.L. Vincent, P.J. Green, D.N. Woolfson, *Bioinformatics* **2013**, *29*, 69.
[8]. T.D. Sutherland, S. Weisman, H.E. Trueman, A. Sriskantha, J.W.H. Trueman, V.S. Haritos, *Mol. Biol. Evol.* **2007**, *24*, 2424.
[9]. T.D. Sutherland, H.E. Trueman, A.A. Walker, S. Weisman, P.M. Campbell, Z. Dong, M.G. Huson, A.L. Woodhead, J.S. Church, *J. Struct. Biol.* **2014**, *186*, 402.
[10]. P.A. Bullough, P.A. Tulloch, *J. Mol. Biol.* **1990**, *215*, 161.
[11]. B.P. Partlow, M. Bagheri, J.L. Harden, D.L. Kaplan, *Biomacromolecules* **2016**, *17*, 3570.
[12]. For further details at a basic level, see: C.E. Housecroft and E.C. Constable, *Chemistry*, 4th ed., Prentice Hall, Harlow, 2010, p. 1321.
[13]. T.D. Sutherland, M.G. Huson, T.D. Rapson, *J. Struct. Biol.* **2018**, *201*, 76.

[a]This column is one of a series designed to attract teachers to topics that link chemistry to Nature and stimulate students by seeing real-life applications of the subject.